



King's Research Portal

DOI:

[10.1093/med/9780198722373.003.0010](https://doi.org/10.1093/med/9780198722373.003.0010)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Sullivan-Bissett, E., Bortolotti, L., Broome, M., & Mameli, M. (2016). Moral and legal implications of the continuity between delusional and non-delusional beliefs. In G. Keil, L. Keuck, & R. Hauswald (Eds.), *Vagueness in Psychiatry* Oxford University Press. <https://doi.org/10.1093/med/9780198722373.003.0010>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Moral and Legal Implications of the Continuity between Delusional and Non-delusional Beliefs

Ema Sullivan-Bissett, Philosophy Department, University of Birmingham

Lisa Bortolotti, Philosophy Department, University of Birmingham

Matthew Broome, Department of Psychiatry, University of Oxford

Matteo Mameli, Philosophy Department, King's College London

Acknowledgements: Lisa Bortolotti, Matteo Mameli and Matthew Broome acknowledge the support of the Wellcome Trust for a project entitled 'Moral Responsibility and Psychopathology' (WT099880MA). Lisa Bortolotti and Ema Sullivan-Bissett acknowledge the support of the Arts and Humanities Research Council for a project entitled 'Epistemic Innocence of Imperfect Cognitions' (AH/K003615/1) and of a European Research Council Consolidator Grant (grant agreement 616358) for a project entitled 'Pragmatic and Epistemic Role of Factually Erroneous Cognitions and Thoughts' (PERFECT). All four authors are grateful to Geert Keil for his helpful comments on an earlier draft of this chapter.

Abstract (248 words)

In this paper we explore two aspects of gradualism as they apply to the phenomenon of delusions, that is, the acknowledgement that it is difficult to distinguish pathological from non-pathological beliefs and the defence of the view that there is considerable continuity between delusional and other epistemically faulty beliefs. We also identify the implications of these two aspects of gradualism for questions about one's moral and legal responsibility for action that is motivated by one's delusional beliefs. In the first section of the paper, we argue that an effective demarcation between pathological and non-pathological beliefs cannot be successfully achieved on mere epistemic grounds, that is, on the basis of the epistemic quality of the relevant belief-forming and belief-maintaining processes. We offer some reasons to endorse the thesis that delusional beliefs are continuous with other epistemically faulty beliefs. In the second section of the paper, we examine the implications of the continuity thesis for the association—common in everyday thinking but also in ethical and

legal frameworks—between being diagnosed with a psychiatric disorder featuring delusions and having reduced or no responsibility for action that is motivated by one’s delusions. We consider some interesting cases of agents who committed crimes related to the content of their epistemically faulty beliefs, and ask whether their beliefs qualifying as delusions makes a difference to whether such agents were responsible for their actions. Finally, we make some suggestions about how the continuity thesis can inform attributions of moral and legal responsibility for action.

Keywords:

Continuity thesis, gradualism, delusion, belief, moral responsibility, legal accountability, epistemic faults.

Introduction

In this paper we explore two aspects of gradualism about mental illness by arguing that it is difficult to distinguish pathological and non-pathological beliefs on the basis of their epistemic features, and examining and ultimately defending the claim that there is no *categorical* difference between delusional and other epistemically faulty beliefs (what we shall call *the continuity thesis*). In section one, we argue that no effective demarcation between pathological and non-pathological beliefs can be achieved on the basis of mere epistemic criteria and we appeal to considerations about the factors influencing belief formation. This seems to support the continuity thesis. In section two, we consider some of the moral and legal implications of the continuity thesis, focusing in particular on the role of epistemically faulty beliefs in the attribution of moral responsibility and legal accountability for criminal actions that are motivated by those beliefs.

1. Delusional and Non-delusional Belief

Belief is an attitude with a standard of correctness, according to which true beliefs are correct and false beliefs are incorrect. We might say that it is ‘part of the “job description”

of belief as a distinctive propositional attitude that beliefs are correct or incorrect depending upon the state of the world' (Railton 1994: 74). While other cognitive states can have contents which are true or false, truth and falsehood are a 'dimension of assessment of beliefs as opposed to many other psychological states or dispositions' (Williams 1970: 136). Correctness conditions then follow not only from the propositional content of a state, but also from the state itself. We also evaluate beliefs with respect to epistemic values other than truth; they are appropriate targets for claims about whether they are *rational* or *justified*. Epistemic norms including norms of evidence ('a belief is correct if it rests upon sufficient evidence'), knowledge ('a belief is correct if and only if it aims at knowledge'), and rationality ('a belief is correct if and only if it is rational') are thought to be ones which govern belief (Engel 2007: 181). These norms govern only belief—it is inappropriate to say of my imaginings or supposings that they are rational, irrational, justified, unjustified, and so on.

Many philosophers have taken such features of belief to highlight something necessary about its nature, and have sought to explain the conditions under which beliefs are formed, and the norms to which we seem to respond. Some philosophers do this by appeal to belief's having an *aim* (McHugh 2011 and 2012; Steglich-Petersen 2006 and 2009; Velleman 2000). Belief, it is suggested, is something which *aims* at the truth, such that, *as believers*, we aim to believe that *p* only if *p* is true.¹ Others have claimed that belief is norm-governed, though there has been considerable debate over what the norm governing belief is. Where normative theorists agree is on the claim that belief is governed by a norm, and it is by appeal to this norm that we can explain why beliefs have a standard of correctness and

¹ At least one other aim of belief has been put forward by Conor McHugh, the aim of knowledge (McHugh 2011). It is beyond the scope of this chapter to discuss the various formulations of the aim account, we only mention it here to make salient the idea that ordinary beliefs are idealised in certain respects. This omission is also acceptable since, as Timothy Chan has pointed out, '[g]iven that knowledge entails truth, if belief aims at knowledge, it also aims at truth' (Chan 2013: 10).

are governed by epistemic norms (see for example Shah 2003; Shah and Velleman 2005; Wedgwood 2002).²

The teleological account explains belief's standard of correctness by pointing out that 'believing p is correct only if p is true because only true beliefs achieve the aim involved with believing' (Steglich-Petersen 2009: 395). The other epistemic norms we highlighted earlier—those of evidence, knowledge, and rationality—are explained by teleological accounts with the claim that 'following them promotes the aim of believing truly' (Steglich-Petersen 2009: 396). If aims have rules or standards associated with the achieving of that aim, then epistemic norms might be considered the rules or standards conducive to achieving belief's aim (McHugh 2011: 371).

According to the normative account of belief, it is part of the *concept* of a belief that a belief is correct if and only if it is true. Other epistemic norms are easily explicable by the normative theorist since according to her account of belief 'the normative properties of belief are constitutive of belief, and are thus explained by the very nature of belief' (Chan 2013: 6).

The aim and norm accounts of belief offer explanations of our epistemic behaviours (such as focusing on the truth when we think about what to believe, gathering evidence, revising beliefs upon new evidence, and so on). The explanations offered involve the claims that belief is constitutively aimed at truth or governed by a norm of truth. It is consistent with such accounts that there can be a break between truth and other epistemic features (there can be a rational false belief, or a justified false belief, for example). But even in cases in which we come to believe something false, we are guided by the aim of belief, or we manifest our commitment to a norm of belief, and these aims or norms are said to be explanatory of our epistemic practice.

Delusions fail to meet many epistemic standards. It might look like they are not beliefs which are aimed at truth or governed by a norm of truth, that they are not responsive

² These are not view that authors endorse, but they are ones which demonstrate how we might think about non-delusional beliefs, their link with truth, and the conditions under which they are formed (see Sullivan-Bissett 2014, chapters two and three for objections to these accounts).

to evidence in the ways which ordinary beliefs typically are. They might be considered as less responsive or even non-responsive to the epistemic norms outlined earlier, which we think other beliefs are responsive to. Differences between delusional and non-delusional beliefs have led some philosophers to argue that delusions are not beliefs at all, but are rather, for example, misidentified imaginings (Currie 2000) or empty speech acts (Berrios 1991).³ The DSM-5 describes delusions as follows:

Fixed beliefs that are not amenable to change in light of conflicting evidence. Their content may include a variety of themes (e.g. persecutory, referential, somatic, religious, grandiose). [...] Delusions are deemed bizarre if they are clearly implausible and not understandable to same-culture peers and do not derive from ordinary life experiences. [...] The distinction between a delusion and a strongly held idea is sometimes difficult to make and depends in part on the degree of conviction with which the belief is held despite clear or reasonable contradictory evidence regarding its veracity. (American Psychiatric Association DSM-5 2013)

As with all definitions of delusions, the DSM-5 definition is controversial, but if we compare it with another influential definition we cannot but notice that the focus is on the epistemic surface features of delusions:

A person is deluded when they have come to hold a particular belief with a degree of firmness that is both utterly unwarranted by the evidence at hand, and that jeopardises their day-to-day functioning. (McKay, Langdon, and Coltheart 2005: 315).

Delusional beliefs are also formed on the basis of insufficient evidence, and may also be incompatible or badly integrated with the persons' other beliefs (Bortolotti and Broome 2008: 822). These characterisations of delusion as *fixed* beliefs, which are *not amenable in light of evidence*, and held *with a degree of firmness that is utterly unwarranted by the evidence at hand*, implies that non-delusional, ordinary beliefs are 'constantly modified by their experiential validation or refutation' (Maher 1988: 32), and so people with delusions are failing to do something which people without delusions routinely do.

³ We will assume a doxastic approach to delusions in this paper (for a defence of doxasticism see Bayne and Pacherie 2005, and Bortolotti 2009, 2012).

Given that aims and norm accounts of belief are seeking to explain the constraints under which people believe (that is, why they focus on evidence, why truth is their guide, and why they are responsive to norms of evidence, knowledge, and rationality in their belief formation), if delusions are not subject to such constraints this might mark them out as different from ordinary beliefs. On the basis of their considerable epistemic faults, delusional beliefs might look different from ordinary, non-delusional beliefs exhibiting a difference in kind and not just in degree. This is precisely the conclusion we seek to resist in this chapter.

1.1. Non-delusional epistemically faulty belief

Here we suggest that non-delusional beliefs are idealised in the psychological and especially in the philosophical literature. We do this by considering two kinds of epistemically faulty belief as they appear in the non-clinical population: beliefs from doxastic biases and beliefs from self-deception. We shall show that these beliefs also exhibit failures of rationality and depart from epistemically ideal practices of belief formation and belief maintenance.

1.1.1. Doxastic biases

A practice is a *doxastic bias* if it is an unreliable (in terms of truth) doxastic practice (Hazlett 2013: 41). The self-enhancement bias is one example of a widespread doxastic bias, and this encompasses ‘overly positive self-evaluation, unrealistic optimism, illusions of control, self-serving causal attributions, valence biases in recall and processing speeds, biased attention to evidence, [and] biased self-focused attention’ (Hazlett 2013: 52). In an oft-cited study which looked into the self-perceptions of people with and without depression, the self-ratings of participants across various dimensions were compared with the ratings of those participants as given by other people. What was found was that the ‘initial self-perceptions of the depressed subjects were less discrepant with observer ratings’ than controls (Lewinsohn et al 1980: 210). The self-ratings of people with depression ‘did not differ significantly’ from those of their observers, whereas controls rated themselves ‘significantly more positively’ than did their observers. People with depression, then, were the ‘most realistic’ with regard to their self-perceptions, whereas controls ‘were engaged in self-enhancing distortions’ (Lewinsohn et al 1980: 211).

Several other studies have shown that most people are vulnerable to positive illusions, considering themselves (and sometimes their romantic partners) to be above average, or better than most other people, when asked about positive traits and abilities. Moreover, people tend to exhibit unrealistic optimism about their future underestimating their likelihood of experiencing negative events, and overestimating their likelihood of experiencing positive events (for a review see Hazlett 2013, and Bortolotti and Antrobus 2015). In the psychological literature, the suggestion has been made that the positive illusions and unrealistic optimism are adaptive and contribute to mental health, making people happier, more productive and creative, more caring, more resilient (Taylor 1989; Sharot 2011). In his discussion of the empirical studies, Peter Railton claims that '[i]t would appear to be part of the normal, healthy operation of one's self image that one discount negative evidence and defy the odds' (Railton 1994: 93).

The biases discussed here serve to modify the standards for sufficient evidence required for belief. People do not treat evidence in the ways that they do on purely epistemic grounds, and non-epistemic factors are involved when people form beliefs about themselves or make predictions about their own future.

1.1.2. Self-deception

In self-deception, beliefs include a motivational element which can involve the misreading or ignoring of evidence in coming to a belief. The motivational element of the belief forming process might be a pro-attitude towards a proposition being *true* (wishful self-deception), a proposition being *believed* (willful self-deception), or a proposition being *false* (dreadful self-deception) (Van Leeuwen 2007: 423–5).

Let us give an example to demonstrate the non-epistemic factors involved in self-deceptive belief formation. Consider the person who has the false and motivated belief that his wife is faithful. There may be evidence available to the person that his wife is unfaithful, insofar as certain features of her behaviour are *perceptually available* to him (he sees that she arrives home late, that she is uninterested in him, and so on). We might think though that the alternative, epistemically more worthy belief that his wife is unfaithful is unavailable in a weaker sense, a kind of *motivational unavailability*. The person, we can presume, is highly

motivated for it to be the case that his wife is faithful (wishful self-deception), or at least, for it to be the case that he *believes* that his wife is faithful (wilful self-deception). Consider another case, a case of a person with anorexia nervosa, who might come to believe that she is overweight, and she is motivated for her being so to be false (dreadful self-deception). Is the person in these cases is trying to aim at the truth when she forms beliefs (as the teleological account of belief would claim)? Or is she responding to a norm of belief (as the normative account of belief would claim)? It might be that she is, but what it is that makes it the case that these beliefs are aimed at truth or governed by a norm of truth might be very different from what is going on in a more epistemically ideal case where motivational factors are not playing a significant role in the fixation of belief.

1.2. Non-epistemic factors in faulty belief

We found common instances of belief that fail to satisfy the standards that delusions also fail to meet. Either the mechanisms responsible for belief-production are not solely geared, in all cases, towards truth, or even when they are, they often miss that target. The cases of belief we have discussed above depart considerably from the idealised conception of beliefs as mental states that are responsive to evidence and revised in the light of counterevidence. The formation and maintenance of the beliefs we considered seem to be paradigmatically influenced by non-epistemic factors.

To further explain what the cases have in common and how they can be regarded as instances of epistemically faulty belief, we can look to Yaacov Trope and Akiva Liberman's (1996) idea of *confidence thresholds* for belief. The idea here is that there is a correlation between a person's confidence threshold and the evidence that is required to reach the threshold, the lower the first, the less evidence required to reach it. The *acceptance threshold* is 'the minimum confidence in the truth of a hypothesis that [one] requires before accepting it, rather than continuing to test it', while the *rejection threshold* is 'the minimum confidence in the untruth of a hypothesis that [one] requires before rejecting it and discontinuing the test' (Trope and Liberman 1996: 253, cited in Mele 2000: 34). What is meant by *cost of information* is the resources and effort that the person requires in order to acquire and process information relevant to the target proposition. What is meant by *cost of false acceptance* and *cost*

of false rejection is the subjective importance the person attaches to avoiding falsely believing a proposition and avoiding falsely believing the negation of a proposition, respectively (Trope and Liberman 1996: 252, cited in Mele 2000: 34). If this model is correct, our desires can influence our beliefs by functioning to change our confidence thresholds: (1) in several cases of doxastic bias, pro-attitudes play a role in belief formation; (2) in the case of self-deception, an attitude towards the target proposition plays a role in generating a belief in that proposition, a belief which would not be acquired were the attitude absent. So belief formation is often influenced by non-epistemic factors, which include motivational factors.⁴

Allan Hazlett suggests that there might be coping mechanisms in the form of self-deception to go some way towards offsetting the negative consequences of bad life events, and that such mechanisms might also give rise to ‘less extreme’ biases which might be ‘useful as means of coping with the events of everyday life’ (Hazlett 2013: 61). Ryan McKay and Daniel Dennett (2009) go as far as to argue for the presence of a doxastic shear pin, a mechanism that allows desires to influence belief formation when the person would be harmed by believing what she has evidence for and would struggle to manage negative emotions. In some of these cases, the epistemically faulty belief (they call it ‘misbelief’) can be biologically or psychologically adaptive. Interestingly, candidates for adaptive misbeliefs include positive illusions and delusions.

We saw that in many cases non-epistemic factors influence the fixation of belief, and this indicates that a different strength of regulation for truth, or responsiveness to evidence, and so on, applies to different instances of believing. Hence it is difficult to group all beliefs together by appeal to their epistemic surface features. To be clear: we are not suggesting that the attitudes resulting from doxastic biases and self-deception are not beliefs, we think that they are. Rather our claim is that it is implausible to suggest that the reason why these cognitions are beliefs is that they share some good epistemic feature with other non-

⁴ It might be that it is even justified to make justification standards and confidence thresholds context-relative. This kind of claim is not the one we are after in this chapter. We are not trying to give a normative account of how beliefs ought to behave, rather, we are doing descriptive work. So we remain neutral on whether it is justified or rational to have lower evidence thresholds in some cases. We are just pointing out that as a matter of fact, we *do* have lower thresholds.

delusional beliefs, and then claim that delusional beliefs are different in kind because they are epistemically poor, or lack some good epistemic feature.

Next we turn to delusions and argue that they are continuous with the epistemic faults detected in the two cases discussed above, namely, doxastic biases and self-deception.

1.3. Delusional belief

Let us turn now to epistemically faulty beliefs that are also delusional. In this section we shall argue for the continuity thesis in two steps. First, we notice how the most popular theories of delusion formation are compatible with or actively support the continuity thesis. Second, we observe that the epistemic faults that characterise delusional beliefs also characterise non-delusional beliefs, and in particular beliefs due to doxastic biases or self-deception.

1.3.1. Delusion formation

Here we cannot provide a detailed description of all the promising theories of delusion formation discussed in the literature, but by appealing to the most influential proposals we aim to show that delusions are best understood as beliefs, and as continuous with non-delusional beliefs. In particular, delusions are seen as understandable (sometimes even rational) responses to anomalous experience. The process by which people form delusions should not be described as radically or categorically different from the process by which people form ordinary beliefs.

According to the one-factor account of delusion formation, people with delusions do not suffer from an *abnormal* deficit or bias in their mechanisms of belief formation or belief evaluation. The difference between a person with delusions and a person without is in the kinds of experiences the person with delusions has. Brendan Maher claimed that ‘delusional beliefs are developed in much the same way that normal beliefs are’ (Maher 1988: 22), and that the experiences people with delusions have are such as to distort the evidence they have available to them. This means that delusions are not held in the face of obvious counterevidence as they are often characterised, but rather are held ‘because of evidence strong enough to support [them]’ (Maher 1974: 99). One-factor accounts do not deny that

reasoning biases might be involved in the process by which people come to delusional beliefs, but claim only that ‘delusions occur when those biases are exaggerated or introduced by intractable anomalous experiences [...] the delusion results from an anomalous experience rationalized by a mind whose divergence from ideal rationality is within the normal range of human psychology’ (Gerrans 2002: 52).

One popular version of the one factor theory is the prediction-error theory proposed by Phil Corlett and colleagues.⁵ When people experience something that does not match with their current understanding of the world, a prediction-error signal is produced and either the input is reinterpreted or the model of the world is revised to take into account the new experience. The hypothesis is that in people with delusions, the production of excessive prediction-error signals falsely suggests that a person’s internal model of the world needs to be updated.

Prediction error theories of delusion formation suggest that under the influence of inappropriate prediction error signal, possibly as a consequence of dopamine dysregulation, events that are insignificant and merely coincident seem to demand attention, feel important and relate to each other in meaningful ways. Delusions ultimately arise as a means of explaining these odd experiences. (Corlett et al. 2009: 1)

On this account, delusion formation differs from the formation of other beliefs only in so far as prediction-error signalling is disrupted. The process of belief formation is the same in the case of delusional and non-delusional beliefs, but the signalling is disrupted in the case of delusions.

According to the two-factor account of delusion formation, we need to appeal to two factors in order to explain why a person comes to form a delusional belief. The first factor is the anomalous experience appealed to by one-factor theorists, but two-factor theorists claim that this is not sufficient for the delusion to be formed or maintained and so some deficit or bias in belief forming or maintaining mechanisms also needs to be posited. Philosophers and psychologists endorsing this view disagree on how to characterise the second factor. Plausible characterisations of the second factor provided so far indicate a

⁵ For another account of delusion formation based on a prediction error model, see Hohwy (2013).

difference in degree, rather than kind, between delusional and non-delusional beliefs. For instance, according to the version of the two-factor theory recently proposed by Max Coltheart and colleagues, people with delusions form beliefs in line with a Bayesian model of abductive inference, according to which ‘one hypothesis H_1 explains observations O better than another hypothesis H_2 just in case $P(O | H_1) > P(O | H_2)$ ’ (Coltheart, Menzies, and Sutton 2010: 271). Considering a case of the Capgras delusion⁶ where a man mistakes his wife for an impostor, the two hypotheses in play are the *stranger hypothesis* (the woman who looks like my wife is not my wife) and the *wife hypothesis* (the woman who looks like my wife is my wife). Coltheart and colleagues argue that ‘the observed data are clearly much more likely under the stranger hypothesis than under the wife hypothesis. It would be highly improbable for the person to have the low autonomic response if the person really was his wife, but very probable indeed if the person were a stranger’ (Coltheart, Menzies, and Sutton 2010: 277)

[I]f the stranger hypothesis explains the observed data much better than the wife hypothesis, the fact that the stranger hypothesis has a lower prior probability than the wife hypothesis can be offset in the calculation of posterior probabilities. And indeed it seems reasonable to suppose that this is precisely the situation with the subject suffering from Capgras delusion. The delusional hypothesis provides a much more convincing explanation of the highly unusual data than the non-delusional hypothesis; and this fact swamps the general implausibility of the delusional hypothesis. (Coltheart, Menzies, and Sutton 2010: 278)

On this view, the second factor explains the maintenance of the delusion. The person does not reject the delusional hypothesis once the disconfirming data starts to come in because he seems to be ‘ignoring or disregarding any new evidence that cannot be explained by the stranger hypothesis. It is as though he is so convinced of the truth of the stranger hypothesis by its explanatory power that his conviction makes him either disregard or reject all evidence that is inconsistent with the hypothesis, or at least cannot be explained by the hypothesis’ (Coltheart, Menzies, and Sutton 2010: 279–80).

⁶ Capgras delusion is the [b]elief that others, often related, have been replaced by identical or near identical others; variations exist in which objects or animals are believed changed; the symptoms may be chronic or permanent’ (Ellis, Luat  , and Retterst  l 1994: 119).

The account of delusion formation proposed by Philippa Garety and David Hemsley (1994) explicitly endorses the continuity thesis. The basic thought is that delusions are formed due to a multiplicity of factors, including past experience, affect, self-esteem, motivation, and biases in reasoning (especially probabilistic reasoning) and perception. Some factors interact with one another, and some are more prominent in the formation of some delusions rather than others. There is no need to hypothesise a radical deviation from normal processes of belief formation and maintenance and some of the biases responsible for the epistemic faults of delusions, such as selective attention, confirmation bias, and jumping to conclusions, may affect people with delusions more than clinical and non-clinical controls, but are not distinctive factors. The multifactorial view acknowledges that many of the biases responsible for the formation of delusional beliefs are biases all people are prone to, and explicitly characterises the difference between delusional and non-delusional beliefs as a difference in degree.

1.3.2. Delusions and other epistemically faulty beliefs

Beliefs formed as a result of doxastic biases are continuous with delusional beliefs as their epistemic faults can be described in terms of the beliefs failing to take into account or respond to statistical evidence that is available to the person. In some cases, people who develop delusional beliefs have the same biases as the rest of the population, but are vulnerable to those biases to a greater extent. For instance, delusional and non-delusional beliefs can be due to the *attribution error*, where the person attributes positive events to herself and negative events to external factors or other people. People who develop persecutory delusions may have an exaggerated tendency to fall prey to the attribution error and other similar biases (Freeman et al. 2002). In other cases, people who develop delusional beliefs have a different bias from the one that affects the rest of the population, but both groups are affected by biases that lead to the formation of epistemically faulty beliefs. For instance, when evaluating evidence for a statement, people tend to wait until they have more clues than they need before coming to a decision. This tendency is often called ‘conservatism’ (Stone and Young 1997; McKay 2012). Empirical evidence suggests that people who develop delusions have the opposite tendency and ‘jump to conclusions’, that is, come to a decision about whether a statement is true without having sufficient evidence (see

Fine et al. 2007, but also Ross et al. 2015). This latter tendency is often called ‘revisionism’. Both tendencies are epistemically problematic, but conservatism is more widespread in the non-clinical population.

Even the epistemic feature of delusions that is considered most distinctive—resistance to counterevidence—is actually a very common feature of beliefs (Bortolotti 2009, chapter 2). Once they adopt a hypothesis, people are very reluctant to abandon it, even when copious and robust evidence against it becomes available. This is true of prejudiced and superstitious beliefs (see, for example, Rusche and Bewster 2008), but also of beliefs in scientific theories (see, for example, Chinn and Brewer 2001), a context in which responsiveness to evidence should be seen as highly important. Self-enhancing beliefs are especially resistant to counterevidence, and people keep believing that they are skilled, talented, attractive, successful, and so on, even when their life experiences repeatedly suggest otherwise. In order to maintain a positive image of themselves, they reinterpret negative feedback and focus on selected evidence that supports their self-enhancing beliefs (Hepper and Sedikides 2012).

Beliefs in the context of self-deception can be vulnerable to a number of doxastic biases and are also resistant to counterevidence. Indeed, non-clinical instances of self-deception have been compared to *motivated* delusions, that is, those delusions that can be construed as playing a defensive function, and other delusions whose formation is affected by what the person desires to be true (McKay and Kinsbourne 2010). Motivated delusions can include erotomania, where a person believes that another is in love with her; grandiose delusions, where the person believes that she is, for example, a largely misunderstood genius; and anosognosia, where the person denies a serious impairment.⁷ In the formation of such delusions, just like in self-deception, motivational influences play a role in the adoption of a belief, and the resulting belief is not well-supported by or responsive to the evidence.

These considerations are, obviously, not conclusive. We have considered how delusional and non-delusional beliefs are formed, and what epistemic faults delusional and non-delusional beliefs are vulnerable to. We might look for the difference between delusions

⁷ See Mele (2008), Davies (2008), and Bortolotti and Mameli (2012) for a discussion of how delusions relate to self-deception.

and otherwise epistemically faulty beliefs elsewhere. Considering how a person reacts when she is made aware of her cognitive biases and confronted with powerful arguments against her belief might introduce a significant difference between delusional and non-delusional beliefs. We might think that a person with delusional beliefs would reject alternative explanations of her beliefs or experiences offhand, whereas a person with non-delusional beliefs would be much more responsive to feedback.

As it happens, the empirical evidence does not support discontinuity in this area. It is well-known that people are very resistant to change their beliefs even when they are told what reasoning mistakes and biases affected the formation of such beliefs (Stalmeier *et al.* 1997; Lichtenstein and Slovic 1971; Tversky and Kahneman 1983), and we already saw that people ignore or reinterpret negative feedback on their own performance to protect self-enhancing beliefs. The claim that people with delusions are resistant to cognitive probing also needs to be qualified. There are strong indications that cognitive behavioural therapy is efficacious in reducing the rigidity of delusional states, and the preoccupation of the person with the topic of the delusion (Coltheart 2005; Kingdon *et al.* 2008). Although the evidence gathered so far does not suggest that cognitive behavioural therapy is effective in leading the person to abandon a delusion altogether, cognitive probing does contribute to the person adopting a more critical attitude towards the content of the delusion (Bortolotti 2009, chapter 2). Thus, focusing on how people respond to challenges is not a promising way to argue for discontinuity between delusional and non-delusional beliefs.

1.4. Interim conclusion

So far we have argued that epistemically faulty delusional and non-delusional beliefs do not differ in kind. Delusions, like other beliefs, are resistant to counterevidence, and the formation of delusions, like the formation of other epistemically faulty beliefs, is influenced by non-epistemic factors. In the case of delusional and non-delusional beliefs alike, there can be considerable resistance to abandoning a belief once it has been adopted and biases and motivational factors may influence belief formation. Next, we move to the moral and legal implications of this view.

2. Moral and legal implications of the continuity view

What factors should be taken into account when attributing criminal responsibility to perpetrators of severe crimes? Here we would like to discuss three cases of people with epistemically faulty beliefs who committed serious offences. Our purpose is to ask whether the presence of *delusional* as opposed to *non-delusional* beliefs is always a reason to doubt the responsibility people have for those actions that seem to be guided or motivated by their beliefs. If there is no categorical difference between delusions and other epistemically faulty beliefs, why is the presence of delusions regarded as a key factor in establishing criminal responsibility?

The first case we consider is that of Bill who attacks a neighbour because he believes the neighbour is shouting insults to him and intends to harm him (Broome, Bortolotti, and Mameli 2010). The second case is that of Jeremiah Wright who killed his son while believing that his son was a CPR dummy (Kotz 2011). The third case is Anders Breivik's perpetration of mass murder in Norway (Bortolotti, Broome, and Mameli 2014).

The analysis of these cases puts some pressure on the view that the presence of delusions is sufficient to determine whether agents are morally responsible and legally accountable for their criminal actions.

2.1. Three cases

2.1.1 Bill

Matthew Broome and colleagues describe the case of a young man with a diagnosis of schizophrenia who attacked his neighbour after experiencing auditory hallucinations about the neighbour making loud noise and insulting him repeatedly. Bill was convicted of assault but his sentence was affected by a pre-existing diagnosis of schizophrenia. He was sentenced to two years' probation and his custodial sentence was suspended.

[S]uppose Bill had actually had a very noisy neighbor. What kind of ascription of responsibility would we have made in relation to the harm inflicted on his neighbor in those circumstances? What kind of punishment would Bill have deserved for his attacking his truly noisy neighbor?

Should the fact that the experiences were hallucinatory (and thereby that the neighbor was not in fact noisy) make a difference in relation to how we conceive of Bill's responsibility for what he did and of the punishment he deserves? It is true that Bill was hallucinating: He was hallucinating that his neighbor was making loud noises, and the content of the hallucination explains in part why he attacked his neighbor. Had he not hallucinated that his neighbor was making loud noises, Bill would have probably not attacked and harmed his neighbor. But it is also true that having noisy neighbors does not morally justify assaulting them. That is, had Bill's neighbor been truly noisy, Bill would have still been doing something blameable in assaulting his neighbor. If one has a noisy neighbor, then one should try to convince his neighbor to be less noisy, and, failing that, one should perhaps call the police. (Broome, Bortolotti, and Mameli 2010: 182)

Here, what we find is that the psychotic symptoms experienced by Bill help *explain* but not necessarily *justify* his aggressive behaviour towards his neighbour. His experiences (auditory hallucinations) and delusional beliefs (the belief that his neighbour intended him harm) help explain why he assaulted his neighbour, but the assault was not inescapable or excusable given such experiences and beliefs.

What we can draw from the case of Bill is that the presence of delusions is not sufficient to regard the person who committed a crime unaccountable due to insanity, though of course the presence of delusions is relevant to the person's full psychological profile at the time the crime was committed and thus should be taken into account. For instance, it is possible that the presence of the delusion signals the presence of reasoning impairments that affect the agent's decision-making capacities.

2.1.2 Jeremiah Wright

Our next case is different from the case of Bill in important ways. On 14th August 2011 Jeremiah Wright killed his seven year old son, Jori, who had cerebral palsy requiring full time care (Kotz 2011). He beheaded and dismembered the child in the home he shared with the child's mother. Wright was charged with, and tried for, first-degree murder. Wright was suffering from a delusion at the time of the killing (as well as before and after the act). He believed that Jori was not his son, but a CPR (Cardiopulmonary Resuscitation) dummy, placed in his home as part of a government experiment. Wright was found not guilty on the grounds of insanity.

A police report stated that 'Wright said that he recently saw the way the dummy looked at him and there were signs and little things the dummy did to him that let him know that Jori was not his son, but a dummy' (Quigley 2013). Dr. Sarah DeLand, director at the mental facility in which Wright was housed, and George Seiden, a psychiatrist working with Wright, testified that Wright believed Jori was a CPR dummy. Wright told DeLand and Seiden that Jori was a government social experiment claiming: 'I don't believe they can do anything to me because it wasn't a real person. His skull was made of plastic. He had foam in him' (Quigley 2013).

Now let us suppose, as we did with the case of Bill, that Wright's beliefs were not delusional, and their content were true. Let us suppose then that Jori, the seven year old boy, was actually a CPR dummy. What ascription of responsibility would we make with respect to the 'harm' inflicted on Jori, and what kind of punishment would Wright deserve? In the case of Bill, his belief that his neighbour was shouting at him would help explain but not justify Bill's assault, as having noisy neighbours does not justify assaulting them. But if Wright had a CPR dummy in his home, then it would not be obviously morally wrong to 'decapitate' and 'dismember' that dummy given that it would not be a living being capable of feeling pain and suffering.

Wright's psychotic symptoms, like Bill's, help explain his behaviour. Bill feels threatened and frustrated because he believes his neighbour is causing him trouble and might intend to harm him. In addition, Bill might think that other courses of action are closed to him given his history of mental illness—calling the police, for instance, may not be an attractive option if Bill suspects that the police will not believe him. Wright wants to prevent the government from spying on him, and thus wants to destroy the dummy. The difference between the two cases is that, in Wright's case, if the content of his belief were true, it would not be morally problematic to destroy the dummy, and the action could be justified by Wright's desire to stop the government's intrusion in his life. Wright's actions would be permitted given his belief that Jori was a CPR dummy. Unlike in the case of Bill then, the presence of Wright's belief is sufficient to regard the person who committed the crime unaccountable, since what Wright did would not be morally problematic if his belief were true. Wright's actions were not *inescapable* as he could have done otherwise given his beliefs, but his delusions seem to offer both some explanation and justification for his actions.

From the first two cases alone it is obvious that the relationship between delusions and criminal responsibility is not a straightforward one. In Bill's case, the delusion went some way toward explaining his action, but it did not justify that action. In Wright's case, his delusions went some way toward explaining and justifying his action as they relieved him of culpability. However, his action was not *inescapable* given his delusional beliefs.

2.1.3 Anders Breivik

In July 2011, Anders Breivik killed seventy-seven people in Norway. In August 2012, he was sentenced to twenty-one years in prison. As part of his first psychiatric evaluation, he was diagnosed with paranoid schizophrenia and some of his most implausible beliefs were regarded as persistent, systematised, and bizarre delusions. For instance, one belief he reported was that he was the leader of a Knights Templar organisation which, according to the Norwegian police, does not actually exist. However, this first assessment leading to the diagnosis of schizophrenia was overruled by a second assessment, according to which Breivik's strange beliefs were not psychotic symptoms in the context of schizophrenia or of some other psychotic disorder, but could be explained by a personality disorder. Based on the fact that he never manifested hallucinations, the second pair of assessors described Breivik's behaviour as caused by a narcissistic personality disorder accompanied by pathological lying (Melle 2013).

If it had been shown that Breivik experienced psychotic symptoms at the time of his crime, then he would have faced trial with a diagnosis of psychosis, and he would have not been regarded as accountable for his actions. This is because, in the Norwegian Criminal Procedure Code, when one has psychotic symptoms, one cannot be attributed criminal responsibility for action: 'a person is not criminally accountable if psychotic, unconscious, or severely mentally retarded at the time of the crime' (Melle 2013: 17). If Breivik's diagnosis of a psychotic disorder had been confirmed, he would have been regarded as 'criminally insane' and sentenced to compulsory psychiatric treatment (Måseide 2012). As a result of the second assessment and his new diagnosis of personality disorder, Breivik was held accountable for his actions as he was thought not to have been psychotic at the time of his criminal act.

Some questions could be raised about the relation of Breivik's beliefs to his actions. Just like Bill could have attempted to talk to his neighbour or call the police instead of planning an assault, and Wright could have removed the 'dummy' from his home, or put it out of sight without destroying it, so too Breivik could have genuinely believed that multiculturalism was one of the greatest harms of Norwegian society without engaging in the actions that led him to kill seventy-seven people. Breivik's thoughts could have been channelled into joining a political party in which such views were shared or campaigning against Muslims and multiculturalism. That is, his beliefs go some way towards explaining his action, but do not justify it and do not make it inescapable.

2.2. Does it matter whether the perpetrator's beliefs are delusional?

The cases we have looked at highlight that we cannot assume that the presence of delusions implies no or reduced responsibility for action. A more local and nuanced view of responsibility needs to be articulated. More precisely, further argument is needed to support the claim that the presence of delusions and other psychotic symptoms is an appropriate criterion for criminal insanity.

In all the three cases we considered (each coming from a different legal jurisdiction, the UK, the US, and Norway), one key question in the psychiatric assessment leading to sentencing was whether the person's system of beliefs was delusional. We saw that the presence of a diagnosis of schizophrenia was instrumental to Bill's lenient sentence. We saw that Wright was found not guilty for reasons of insanity and was committed to a psychiatric hospital for care. And we saw that the presence of delusions alone, if confirmed by the second psychiatric assessment, would have indicated Breivik's lack of responsibility for his mass murder for the Norwegian Law.

The continuity thesis we have defended in section one makes it problematic to rely so heavily on the presence of beliefs that are delusional when assessing responsibility. For claims about responsibility, the significance of the presence of delusional beliefs may derive from the following consideration. If poor reality testing (or some other relevant cognitive deficit associated with delusion formation) is affecting the beliefs a person is prepared to endorse to the extent that such beliefs are implausible even to members of the person's same

culture or subculture, then maybe such failure of reality testing (or other relevant cognitive deficit) is also implicated in some of the person's decision making processes, including those processes that led the person to acting criminally. But this link between the presence of psychotic symptoms and impaired decision-making is just a hypothesis that needs to be tested.

The assumption that people who have psychotic symptoms or have received a diagnosis of schizophrenia lack responsibility or have reduced responsibility for action because their decision-making capacities are impaired is especially problematic, as the behaviour of two people with psychosis or schizophrenia can differ almost entirely. Some people with schizophrenia are able to function well, cognitively and socially, and to control their delusions to some extent. The presence of psychiatric symptoms and of a diagnosis of schizophrenia should be taken into account in the courtroom, but it should not be regarded as sufficient to determine responsibility.

3. Conclusions and implications

In section one we defended gradualism with respect to the distinction between delusional and non-delusional epistemically faulty beliefs. We argued that there is continuity between them: they can be resistant to counterevidence and their formation process may be influenced by biases and motivational factors. Reflecting on the recent psychological literature on delusions, we saw that the mechanisms posited to explain the adoption of delusional hypotheses are not radically different from, but continuous with, standard mechanisms of belief formation.

In section two we turned to the implications of the continuity thesis for moral and legal issues concerning responsibility for action. How should we view the presence of delusions, which is often considered as a key criterion for criminal insanity, if there is no clear demarcation between delusional and non-delusional beliefs to be made on epistemic grounds? We argued that the role of delusional beliefs in motivating action does not seem to be different from the role of other epistemically faulty non-delusional beliefs, unless we assume that the presence of delusions also signals the presence of a cognitive deficit that impacts on the decision to commit the crime in question.

Moreover, we suggested that having beliefs that are epistemically faulty, whether delusional or not, rarely provides a justification for criminal action. It may contribute to an explanation for the crime, but in most cases it does not make the criminal action inescapable or excusable.

References

- American Psychiatric Association 2013: *Diagnostic Statistical Manual of Mental Disorders*, Fifth edition (DSM-5), Washington, DC.
- Bayne, Tim and Pacherie, Elisabeth 2005: 'In Defence of the Doxastic Conception of Delusions'. *Mind and Language*. Vol. 20, no. 2, pp. 163–88.
- Berrios, G.E 1991: 'Delusions as 'Wrong Beliefs: A Conceptual History'. *British Journal of Psychiatry*. Vol. 159, pp. 6–13.
- Bortolotti, Lisa 2009: *Delusions and Other Irrational Beliefs*. UK: Oxford University Press.
- Bortolotti, Lisa 2012: 'In Defence of Modest Doxasticism about Delusion'. *Neuroethics*. Vol. 5, pp. 39–53.
- Bortolotti, Lisa, and Broome, Matthew R. 2008: 'Delusional Beliefs and Reason Giving'. *Philosophical Psychology*. Vol. 21, no. 6, pp. 821–41.
- Bortolotti, Lisa, Broome, Matthew, and Mameli, Matteo 2014: 'Delusions and Responsibility for Action: Insights from the Breivik case'. *Neuroethics*. Vol. 7, pp. 377–82.
- Bortolotti, Lisa and Mameli, Matteo: 2012: 'Self-deception, Delusion and the Boundaries of Folk Psychology'. *Humana.Mente*. Vol. 20, pp. 203–21.
- Bortolotti, Lisa, and Antrobus, Magdalena 2015: 'Costs and benefits of realism and optimism'. *Current Opinion in Psychiatry*. Vol. 28, no. 2, pp. 194–198.
- Broome, Matthew, Bortolotti, Lisa, and Mameli, Matteo 2010: 'Moral Responsibility and Mental Illness: A Case Study'. *Cambridge Quarterly of Healthcare Ethics*. Vol. 19, no. pp. 179–87.
- Chan, Timothy 2013: 'Introduction: Aiming at Truth'. In Chan, T (ed.) *The Aim of Belief*. USA: Oxford University Press, pp. 1–16.
- Chinn, Clark A., & Brewer, William F. 2001: 'Models of Data: A Theory of How People Evaluate Data'. *Cognition and Instruction*. Vol. 19, no. 3, pp. 323–93.
- Coltheart, Max 2005: 'Delusional belief'. *Australian Journal of Psychology*. Vol. 57, no. 2, pp. 72–76.
- Coltheart, Max, Menzies, Peter, and Sutton, John 2010: 'Abductive Inference and Delusional Belief'. *Cognitive Neuropsychiatry*. Vol. 15, no. 1, pp. 261–87.
- Corlett, Philip R., Krystal, John H., Taylor, Jane R., and Fletcher, Paul C. 2009: 'Why Do Delusions Persist?' *Frontiers in Human Neuroscience*. Vol. 3, no. 12, pp. 1–9.
- Currie, Gregory 2000: 'Imagination, Delusion and Hallucinations'. *Mind and Language*. Vol.

15, no.1, pp. 168–83.

Davies Martin 2008: 'Delusion and Motivationally Biased Belief: Self-Deception in the Two-Factor Framework'. In: Bayne Tim, Fernández Jordi, editors. *Delusions and Self-Deception: Affective Influences on Belief Formation*. Psychology Press; Hove: pp. 71–86.

Ellis, H. D., Luauté, J. P., and Retterstøl, N. 1994: 'Delusional Misidentification Syndromes'. *Psychopathology* Vol. 27, pp. 117–20.

Engel, Pascal 2007: 'Belief and Normativity'. *Disputatio*. Vol. 2, no. 23, pp. 179–203.

Fine, Cordelia, Gardner, Mark, Craigie, Jillian, and Gold, Ian 2007: 'Hopping, Skipping or Jumping to Conclusions? Clarifying the Role of the JTC Bias in Delusions'. *Cognitive Neuropsychiatry*. Vol. 12, no. 1, pp. 46–77.

Freeman, Daniel, Garety, Philippa A., Kuipers, Elizabeth, Fowler, David, Bebbington, Paul E. 2002: 'A Cognitive Model of Persecutory Delusions'. *British Journal of Clinical Psychology*. Vol. 41, pp. 331–47.

Garety, Philippa A., and Hemsley, David 1994: *Delusions. Investigations into the Psychology of Delusional Reasoning*. Oxford: Oxford University Press.

Gerrans, Philip 2002: 'A One-Stage Explanation of the Cotard Delusion'. *Philosophy, Psychiatry, & Psychology*. Vol. 9, no. 1, pp. 47–53.

Hazlett, Allan 2013: *A Luxury of the Understanding: On the Value of True Belief*. Oxford: Oxford University Press.

Hepper, Erika G., & Sedikides, Constantine 2012. 'Self-enhancing feedback'. In R. Sutton, M. Hornsey and K. Douglas (Eds.), *Feedback: The Handbook of Praise, Criticism, and Advice* (Chapter 4, pp. 43-56). London: Peter Lang.

Hohwy, Jakob 2013: *The Predictive Mind*. USA: Oxford University Press.

Kingdon, D., Ashcroft, K., Turkington, D. 2008: 'Cognitive behavioural therapy for persecutory delusions: three case examples'. In D. Freeman, R. Bentall, and P. Garety (Eds.) *Persecutory Delusions. Assessment, Theory and Treatment*. Oxford: Oxford University Press; 393–410.

Kotz, Pete 2011: 'Jori Lirette, 7, Dismembered By Jeremiah Wright, His Mom's Boyfriend'. *True Crime Report*. 18th August 2011.
http://www.truecrimereport.com/2011/08/jori_lirette_7_dismembered_by.php

Lewinsohn, Peter M., Mischel, Walter, Chaplin, William, and Barton, Russell 1980: 'Self Competence and Depression: The Role of Illusory Self-Perceptions'. *Journal of Abnormal Psychology*. Vol. 89, no. 2, pp. 203–12.

Lichtenstein, S. and Slovic, P. 1971: Reversals of preference between bids and choices in

gambling decisions. *Journal of Experimental Psychology* Vol. 89, no. 1, pp. 46–55.

Maher, Brendan 1974: 'Delusional Thinking and Perceptual Disorder'. *Journal of Individual Psychology*. Vol. 30, no. 1, pp. 98–113.

Maher, Brendan 1988: 'Anomalous Experience and Delusional Thinking: The Logic of Explanations'. In Oltmanns, Thomas and Maher, Brendan (eds.) *Delusional Beliefs*. USA: John Wiley and Sons.

Måseide, Per Helge 2012: 'The Battle about Breivik's Mind'. *The Lancet*. Vol. 379, p. 2413.

McHugh, Conor 2011: 'What Do We Aim At When We Believe?' *Dialectica*. Vol 65, no. 3, pp. 369–92.

McHugh, Conor 2012: 'Belief and Aims'. *Philosophical Studies*. Vol. 160, pp. 425–39.

Mele, Alfred 2000: *Self-Deception Unmasked*. USA: Princeton University Press.

Mele Alfred 2008: 'Self-deception and Delusion'. In: Bayne T, Fernandez J, editors. *Delusions and Self-Deception: Affective Influences on Belief Formation*. Psychology Press; Hove: pp. 55–70.

Melle, Ingrid 2013: 'The Breivik Case and What Psychiatrists Can Learn from it'. *World Psychiatry*. Vol. 12, no. 1, pp. 16–21.

McKay, Ryan 2012: 'Delusional inference'. *Mind & Language*. Vol. 27, no. 3, pp. 330–355.

McKay, Ryan, Langdon, Robyn, and Coltheart, Max 2005: "'Sleights of Mind": Delusions, Defences, and Self-Deception'. *Cognitive Neuropsychiatry*. Vol. 10, no. 4, pp. 305–26.

McKay, Ryan and Dennett, Daniel 2009: 'The Evolution of Misbelief'. *Behavioral and Brain Sciences*. Vol. 32, pp. 492–510.

McKay, Ryan and Kinsbourne, Marcel: 'Confabulation, Delusion, and Anosognosia: Motivational Factors and False Claims'. *Cognitive Neuropsychiatry*. Vol. 15, no. 1–3, pp. 288–318.

Ross, Robert Malcolm, McKay, Ryan, Coltheart, Max, and Langdon, Robyn 2015: 'Jumping to Conclusions About the Beads Task?' A Meta-analysis of Delusional Ideation and Data-Gathering'. *Schizophrenia Bulletin*. doi: 10.1093/schbul/sbu187

Rusche, Sarah E. and Brewster, Zachary W. 2008: "'Because They Tip for Shit!": The Social Psychology of Everyday Racism in Restaurants'. *Sociology Compass*. Vol. 2, no. 6, pp. 2008–29.

Quigley, Rachel 2013: 'Father "who hacked his disabled son's head off and left it by the road for his mom to see believed the boy, 7, was a dummy or robot"'. *The Daily Mail*. 31st January 2013 (updated version).

<http://www.dailymail.co.uk/news/article-2271394/Jeremiah-Wright-Father-hacked-disabled->

[sons-head-left-road-mom-believed-boy-7-dummy-robot.html](#)

Railton, Peter 1994: 'Truth, Reason, and the Regulation of Belief'. *Philosophical Issues*. Vol. 5, pp. 71–93.

Shah, Nishi 2003: 'How Truth Governs Belief'. *The Philosophical Review*. Vol. 112, no. 4, pp. 447–82.

Shah, Nishi and Velleman, J. David 2005: 'Doxastic Deliberation'. *The Philosophical Review*, Vol. 114, no. 4, pp. 497–534.

Sharot, Tali 2011: *The Optimism Bias: A Tour of the Irrationally Positive Brain*. USA, New York: Pantheon.

Stalmeier, Peep F. M., Wakker, Peter P. and Bezembinder, Thom G. (1997): 'Preference reversals: violations of unidimensional procedure invariance'. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 23, no. 4, pp. 1196–1205.

Steglich–Petersen, Asbjørn 2006: 'No Norm Needed: On the Aim of Belief'. *The Philosophical Quarterly*, Vol. 56, no. 225, pp. 499–516.

Steglich–Petersen, Asbjørn 2009: 'Weighing the Aim of Belief'. *Philosophical Studies*, Vol. 145, no. 3, pp. 395–405.

Stone, Tony and Young, Andrew W. 1997: 'Delusions and Brain Injury: The Philosophy and Psychology of Belief'. *Mind & Language*. Vol. 2, no. 3–4, pp. 327–64.

Sullivan-Bissett, Ema 2014: 'Belief, Truth, and Biological Function'. PhD Thesis, University of York.

Taylor, Shelley E. 1989: *Positive Illusions: Creative Self-Deception and the Healthy Mind*. USA, New York: Basic Books.

Trope, Yaacov and Liberman, Akiva 1996: 'Social Hypothesis Testing: Cognitive and Motivational Mechanisms'. In Higgins, Edward and Kruglanski, Arie (eds.) *Social Psychology: Handbook of Basic Principles*. New York: Guildford Press, pp. 239–70.

Tversky, A. and Kahneman, D. 1983: 'Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment'. *Psychological Review*. Vol. 90, no. 4, pp. 293–315.

Van Leeuwen, D. S. Neil 2007: 'The Product of Self-Deception'. *Erkenntnis*. Vol. 67, pp. 419–37.

Velleman, David J. 2000: *The Possibility of Practical Reason*. Oxford: Oxford University Press.

Wedgwood, Ralph 2002: 'The Aim of Belief'. *Philosophical Perspectives*, Vol. 16, pp. 267–97.

Williams, Bernard 1970: 'Deciding to Believe', in his *Problems of the Self*. Cambridge: Cambridge University press, pp. 136–51.